DOCUMENT RESUME

ED 118 206 JC 750 135

AUTHOR Babski, Carl
TITLE Does Students' Expectation of Teachers Affect Students' Evaluation of Teachers?
NOTE 40p.; Practicum for Ed.D., Nova University

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage
DESCRIPTORS Bibliographies; *Evaluation Criteria; Junior Colleges; Learning Experience; Literature Reviews; *Student Attitudes; Student Characteristics; Student Evaluation; *Student Teacher Relationship; *Teacher Evaluation; *Teacher Role; Teaching Methods

ABSTRACT

This report gives an extensive review of the literature dealing with student evaluation of faculty, and investigates the effect of a previously unexplored variable, students' expectations of the teaching-learning situation. Eight student perceptions of the teaching-learning situation were identified: dogmatic, erotic, moral, therapeutic, intellectual, existential, humanistic, and pragmatic. These descriptions of student perceptions were worded into statements for the students to rank on a five-point Likert scale, according to their perceptions of what teaching and learning roles mean. This survey instrument was administered to classes in all four divisions at Miami-Dade Community College, along with questions designed to evaluate the course instructors, and questions intended to elicit information about the students answering the survey. Findings showed that there was no significant difference in the evaluation results among students with different perceptions of teaching and learning roles. The survey instrument is appended, as is a bibliography of literature on the subject of student evaluation of faculty. (NHM)

**********
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
**********

ED118206

# DOES STUDENTS' EXPECTATION OF TEACHERS AFFECT STUDENTS' EVALUATION OF TEACHERS?

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

by

Carl Babski, M.S.

MIAMI-DADE COMMUNITY COLLEGE

A PRACTICUM FOR NOVA UNIVERSITY IN
PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF DOCTOR OF EDUCATION

NOVA UNIVERSITY

JC 760 135

ABSTRACT

This project studied one variable which affects students' responses to teacher behavior. That variable which seemingly has not been intensively studied answered the question "Do students' expectations of the teaching-learning situation or student perception of teaching have any effect upon their responses?"

Selected statements that reflected the students' perceptions of the learning-teacher situation were used to see if differences in orientation or perceptions differ in student evaluation of teacher behavior. Findings show that students' perception as associated with differences in student evaluation does not differ significantly from that which could be attributed to mere chance fluctuations within our total population. A review of the research finding on reliability, validity, and other aspects of student evaluation is presented.

ACKNOWLEDGEMENT

The author wishes to thank Gus Wenzel, John Scirba, and Fred Smith of the Testing and Research Center at Miami-Dade Community College, North Campus, for their guidance and patience.

TABLE OF CONTENTS


| | PAGE |
|---|---|
| LIST OF TABLES | v |
| INTRODUCTION | 1 |
| BACKGROUND AND SIGNIFICANCE | 3 |
| Common Criticism | 3 |
| Research Findings | 5 |
| Effectivness of Student Feedback | 6 |
| Teacher Characteristics | 7 |
| Reliability | 7 |
| Validity | 8 |
| Sex, Major and Other Variables | 9 |
| Student Perception and Value System | 10 |
| PROCEDURE | 12 |
| Population | 15 |
| Procedure for Treating Data | 15 |
| Limitations | 17 |
| RESULTS | 17 |
| Characteristics of the Population | 17 |
| Significance of Data | 18 |
| Analysis of Variance | 18 |
| Internal Reliability | 19 |
| CONCLUSIONS AND RECOMMENDATIONS | 19 |

TABLE OF CONTENTS (cont'd)

PAGE

BIBLIOGRAPHY..... 31

APPENDICES..... 22

A. Summary of arguments for student evaluation.. 22

B. Student evaluation of the Teaching of Miami-Dade course..... 26

C. Sample characteristics by percentages..... 30

LIST OF TABLES


| TABLE | | PAGE |
|---|---|---|
| 1. | Means and Standard Deviations..... | 18 |
| 2. | Analysis of Variances..... | 18 |

## INTRODUCTION

Today in a time of course relevance, legislative accountability, and faculty retrenchment, the students are demanding a greater participatory place for themselves. The Zeitgeist is perhaps helping formulate such student need for a "say so" in shaping the college future, and at Miami-Dade Community College the voice of the student is being heard. The student government has a place in the faculty senate, there is a student representative on the Board of Trustees and now the students, as in many colleges, wish to have their opinions felt in the area of faculty behavior in the classroom.

The Director of Natural Sciences, and the new Director of Arts and Sciences, have been directed to address their Division to a study of teacher behavior in the classroom. A larger explanation of the idea is needed. What the new director wishes or foresees is a feedback system from students for teachers seeking to improve learning by student to teacher behavior feedback and course feedback for improving the course.

To fulfill that directive, a committee consisting of the Chairperson for Physics, two other professors, and myself have been charged to instigate a study of student evaluation of teacher behavior in the classroom as a vehicle to improve learning. This will be the first step in succeeding steps to develop a system of student feedback on teacher behavior and course evaluation.

The discussion at the first meeting entered into many aspects of feedback systems of student evaluation of teacher behavior. As a result of the initial meeting, my assignment is to investigate the common

criticisms of student evaluation of teacher behavior as found in the literature or research.

In association with the work on this committee, it is my feeling that in the vast research concerning the variables which affects student responses to teacher behavior, there seems to be one element which has not been extensively studied. Do students' expectations of the teaching-learning situation or student perception of teaching have any effect upon their responses? The specific purpose of my practicum will be to investigate this variable and to report my results to the committee. This study of student perception of teaching and the review of the literature will constitute my complete report to the committee and my practicum.

This practicum has additional college-wide significance as it is the first developing step in a program to improve learning through student-to-teacher feedback of teacher behavior and the courses.

## BACK GROUND AND SIGNFICANCE

Teaching competence is not easily documented. The teacher is one of the few professionals whose work is seldom observed by his peers. His teaching ability is often based more on hearsay than on substantial evidence. For this reason, promotion committees frequently make final tenure decisions without seriously considering information about teaching. To counter this tendency, the Organizational Review Steering Committee at Miami-Dade Community College recommended to the Vice-President at the North Campus a system of student instructional ratings and, as previously stated, Division directors have formed committees to study this idea. Unfortunately, the literature which will be presented reveals that the faculty have very little more credibility with student evaluation than the hearsay system.

### Common Criticism

Elbe (1974) in his year study of "The Recognition and Evaluation of Teaching" gathered criticisms from an Advisory Board consisting of eleven members from different universities. Their opinions formed from their research is summarized below.

1. Student evaluation does little general good and some particular harm. Faculty anxieties on this point are not precisely clear, for the same critics who are most dubious about general impact are often very sensitive to the particular harms student ratings can inflict. Probably evaluation has impact, good and bad, in relation to the degree it moves away from the strictly private and personal. The effects of published evaluation are not so different from the effects of publishing books, speaking out in public, committing oneself in any way beyond the confines of one's own study.

2. Student evaluation may arouse an unhealthy competition among faculty members.

3. Evaluation systems move toward harmful formalization, mechanization, rigidity.

4. Student opinion questionnaires furnish inadequate or misleading information about teachers and teaching.

5. The effect of student evaluation is a short term one connected with the novelty of the procedure.

6. Students by themselves will not be able to sustain a high level of evaluation procedures over a period of time.

Klierman (1975) lists eleven most provocative items which reflect faculty attitudes. The most notable are summarized below:

1. Student evaluations are used when they support the prior affirmative or negative decision of the college committee or administration.

2. All students and faculty interviewed by Dr. Irene R. Kierman indicated that beginning students cannot usually judge breadth or depth of a teacher's knowledge of his subject. Yet, despite this, student evaluation forms ask this question.

3. In evaluation, what percentage of students should be discounted to compensate for those students who are hostile toward teachers in general?

4. Are we not allowing students to make unsubstantiated and anonymous charges via unsigned student evaluation?

5. How does subject matter which a teacher must handle affect the evaluation he receives?

6. How can we account for racial, ethnic, and sex bias of students in evaluating teachers?

7. How valid are our student evaluations?

Ten of Kierman's questions deal with political, sociological, and psychological aspects of student evaluation currently in use. The last question concerns a legal matter, one of using invalidated tests which may have the effect of discriminating against women and minority group members in matters of employment. Assuming that student evaluation forms in use have not been validated, they do not comply with Equal Opportunity Commission regulations. (Higher Education Guidelines, Executive Order 11246, Page J6)

Research Findings

The primary justification for the heavy reliance on student ratings uses the argument that the student, as the primary consumer of the teaching product, is the best positioned person to evaluate its worth. Roden (1973) studies produced a -.75 correlation between the student evaluation and how much the student learned from the teacher.

Frey (1973) had some doubt about the significance of this negative correlation because it involved teaching assistants. However, Frey's (1973) dubious attitude toward Roden was strengthened by a research report by Arthur Sullivan at the Memorial University of Newfoundland. Sullivan involved ten courses of thirty students each, all classes meeting at the same time. In each class students used a common syllabus, a common text, and took a common final exam. All exams were scored by a special committee. The results indicated a positive relationship ($r = .3$ to a mean $r = .39$ for the ten classes) between student ratings of teachers and external criteria of teaching performance.

Some years ago, Cohen and Brawer (1969) insisted that student gains toward achievement of specific objectives of a course are the "ultimate" criteria for evaluating effective faculty performance. Cronback and Furly (1970) encounter with measuring students gains as an index of "good teaching" is fraught with practical and technical difficulties.

Two studies were completed at Purdue University comparing the amount learned from chemistry instructors with student ratings of the instructors. The material covered and the grading procedure were more or less standardized. An effort was made to correct for the effect of difference between classes in initial ability. In both of these studies, a low positive correlation was obtained between how a class rated the instructor and how much material they had learned (about +.24). Roden (1973) in his articles, concluded that "if how much students learn is considered to be a major component of good teaching, it must be concluded that good teaching is not validly measured by student evaluation in their current form."

Effectiveness of Student Feedback. Centra (1972) undertook a study in 1971-72 with five different type colleges to test if feedback to teachers from student evaluation would improve teaching. This very extensive research divided teachers into three testing groups of feedback, no feedback, and a control group. The results proved no significant changes among them. Included in the student ratings were items that faculty members in an earlier study had identified as information they

they would like to obtain from students (Centra 1972 ). The surprising results were that nearly all three groups were nearly identical in their scores, an indication that the group of instructors who received student feedback did not noticeably modify their teaching practices. However, after the research was extended, the time factor did produce a change; but individual counseling also helped. Centra commented,"The changes were by no means overwhelming, but it does support the utility of student ratings for instructional improvement." However, the evidence is very meager and not of a magnitude to encourage using student evaluation unless an extensive counseling system or what some may label "watch dog system" is utilized to promote improvement. Eble (1971) suggests making the ratings public and this would bring better results.

Teacher Characteristics. Brawer (1968) reviewed studies of the personality characteristics of college and university faculty and found that the studies were "few and inconclusive." Cohen and Brawer (1969), studying the characteristics that a successful teacher would possess, discovered that the teacher's characteristics are highly diverse. Research findings to date have not been especially fruitful.

Reliability. Frey (1973) demonstrated that students can agree about the strength and weaknesses of different teachers. Costin, Greenough and Menges (1971) agree that students can rate classroom instruction with a degree of reliability. This conclusion was reached after extensive review of the findings concerning reliability. It must be stated that these studies were at a correlation of .2 to .3 in many areas.

Validity. Even though students are the only people who observe a teacher's course daily, many faculty seriously question the validity of students' instructional ratings. Some faculty claim to teach philosophical values or to inculcate a special attitude toward learning rather than simply help students to master the subject matter. These goals are difficult to measure and, therefore, it is difficult to determine whether teachers who successfully accomplish these goals are also the ones who are rated highly by their students.

Many researches can be listed which might culminate into validity of student rating under some criteria headings such as:

1. Consumer Satisfaction and Course Objectives. McKeachie (1969 ) found that students who performed particularly well on the test rated their teachers as "more effective" than did students who performed more poorly on the test. McKeachie, Lin and Mann (1971 ) reported similar findings and students who rated teachers high on "rapport" (i.e., warmth) tended to be effective on measures of students' "critical thinking." Lathrop and Richmond (1967) found positive correlation between high ratings and the extent to which students thought they had achieved course objectives.
2. Students' Criteria of Effective Teaching. Crawford and Bradshaw (1968) asked students to describe the most effective teacher they had ever had. The most frequent characteristics

were: (a) knowledge of subject (b) planned and organized lecture (c) interest in teaching (d) warmth toward students. French (1957) had a similar test as did Elliott (1950), Mann (1950), and McKeachie (1969b)

Another point of interest is to relate students' ratings to the degree of teachers' experiences or training. Walker (1969) found that students' ratings tended to improve with the experience of the faculty member. A full professor tended also to receive higher student rating than did other ranks (Downie 1952), especially on certain items as sense of humor, presentation, and interest.

Sex, Major, and other Variables. Costein (1971) implies to us that sex does enter into student evaluation.

1. In four of the five studies teachers rated high on "Skill" tended to be effective with women students.

2. In all five studies teachers rated high in "Structure" tended to be more effective with women than with men. In fact, on the whole, the more structured instructors tend to be ineffective for male students.

3. Teachers who were high in "Rapport" ("warmth") tended to be effective on measures of student thinking.

4. Teachers whom students rated as having an impact on beliefs were effective in changing attitudes.

Majors in psychology were found to rate courses and instructors about the same as non-majors. However, students required to take a psychology course tended to rate it lower than did students who selected course (Cohen and Humphreys 1960). Other investigators found (Gage 1961, Lovell and Haner 1955) that teachers of required courses

received significantly lower student ratings than did teachers of elective courses.

Non-proponents of students' ratings frequently suggest that teachers of large classes may receive lower ratings because students prefer more interaction which occurs in smaller classes. The list of investigators that support this is extremely long. The most recent investigation of this belief is McDaniel and Feldhusen, 1970.

Another argument about the validity of student rating is that students may judge instruction on the basis of its "popularity" value. Guthrie (1954) concluded that "Popularity" of a teacher may well be an index of "good and substantial" teaching and students' ratings may reflect both.

Student Perception and Value Systems. In the concluding discussion of a research by McKeachie, Lin, and Mann (1971), the following proposition was concluded:

> So student ratings have some usefulness. Why aren't they better? Our best guess is that the major slippage in our validity studies is in the differing goals of teachers and students. Students come to a class with many different personal objectives for that class. Some of these objectives may coincide with those of the instructor, but the overlap between instructor and student goals or between the goals of differing students is certainly far from perfect. The extent to which our tests measure achievement of these goals is also limited. Thus, even if each student's rating of the instructor's effectiveness in helping that student achieve his own goals were perfectly valid, we would find low validity coefficients in studies such as this.

Frey (1974) had this to say:

> Thirdly, I believe that it is important to take into account the fact that student's perceptions are a product of their own personalities as well as of the teacher's behavior. Thus the impression that a teacher creates depends not only on his own behavior but also on the behavior and expectations of his audience. Any analysis which assumes that teacher ratings are independent of their source will be woefully inadequate.
>
> Evaluation for each individual, be he student, faculty, or administrator, consists of taking what the teacher said and did and considering these factors in terms of one's own value system and educational philosophy. Certainly, one student may value one aspect of the course more than a second student would. These considerations imply that any overall rating that a student makes will reflect two different types of factors; the events which the student observed during the course and his evaluation of these events within his own value system.

Frey (1974) using the term "in his own value system" is getting very close to my agreement with Epperson's (1974) orientations or expectations toward the teaching-learning situation. McKeachie's "goals of the students" is more loosely interpreted as perceptions, but goals are effected by our value system and goals, value systems and perceptions are all interrelated in the total complex being.

Frey and McKeachie are relating to my research that student perception, which is part of the student's value system, may have effects on the responses to student evaluation. Frey, in his discussion, suggested that students can provide teacher evaluation by constructing the proper questions. A rating form can be constructed with appropriate questions that are uncontaminated by students'

value system. Frey is presently developing such a form.

Committee Report. Because of sickness of committee members the final report was not completed. This resulted in myself being selected to summarize what materials the members had in their possession. The summary is in Appendix A.

## PROCEDURE

David C. Epperson in his article, "Assessing Alternative Teaching-Learning Alliances" in L. J. Stiles, Ed., Theories of Teaching, (New York: Dodd, Mead and Co., 1974) Ch. 5., suggests students may have at least eight orientations or expectations toward the teaching-learning situation. He listed these as:

1. Dogmatic - The teacher is an authority in his field and teaches the course with little room for differences of opinion.
2. Erotic - The teacher's role is to develop a climate in which the student may express his inner feelings and clarify them to himself.
3. Moral - The teacher's major role is to point out the gap between what ought to be done in today's society about social problems and what is being done about them. Hopefully, the student, will be motivated to actively participate in the problem-solving activities.

4. Therapeutic - The teacher's role is to help the student resolve the conflicts in society, such as justice vs, injustice; equality vs. inequality.

5. Intellectual - The course tackles tough, complex intellectual problems.

6. Existential - The student is happiest in courses in which he finds direction and purpose for his life.

7. Humanistic - Classes in which the student's uniqueness and individuality are highlighted are most helpful.

8. Pragmatic - Those courses in which practical, saleable skills are developed are most enjoyable to the students.

Epperson's description of these perceptions were worded into statements for the student to rank on a Likert scale which accurately reflects his concept of what teaching and learning roles mean.*

The developed instrument that was used in this practicum is in Appendix B attached. The statements for Epperson's descriptions were developed and reviewed by an expertise group. Three statements for each of the Epperson descriptions were developed. Then my expertise group of four faculty members selected that statement which they considered to best fit the description. The expertise group consisted

* The reader must read the directions given to the student. The student only selected one best teaching perception with the rank of "strongly agree."

of one English teacher, one from testing, one from physics, and one from the faculty development division. Suggested revisions were reviewed and the selected statements were finalized. My selected statements thus reflect the descriptions of Epperson as judged by my expertise group. These eight orientations or expectations toward the learning-teacher situation will be evaluated to see if differences in orentiation or expectations are associated with differences in student evaluation of teacher behavior.

The evaluation form will be composed of three parts:

| | |
|---|---|
| Part I: | Perception of Teaching Statements |
| Part II: | Teacher Behavior Evaluation |
| Part III: | Control Data |

Part I was prepared by me and reviewed by a board of expertise composed of four faculty members. Part II was coordinated with Organization Review Steering Committee and the Research and Testing Department. The ORSC had introduced student evaluation as a discussion subject at a recent workshop for some fifty faculty members. From that workshop was developed a pilot evaluation form. After some review and discussion with Dr. John Losak, MDCC, and Dr. John Alteman of FIU, my decision was to use the Student Evaluation of the Teaching of Miami-Dade Course Form developed at Miami-Dade north. This evaluation form is an instrument which has been used for a number of years. Part III includes additional information which may be utilized for other significant results. Discussions of the control data were also conducted with

the above mentioned individuals and members of the committee and faculty.

Correspondence has been initiated with Inter-institutional Research Council of Florida Community Colleges and SIR, the Student Instructional Report, of the Educational Testing Service of New Jersey. They have numerous studies and systems which need the committee's attention.

Population. The population to which the study is intended is the student body at Miami-Dade Community College (North). Since the North Campus is divided into four divisions, my idea was to randomly select some teachers from each of the four divisions. This selection from a cross section of four divisions was intended to give a sample more adequate than random selection from the total faculty because randomized selection of teachers might result in selection of the entire sample from the same division. Once the teachers had been selected they would, by random choice of classes, administer the questionnaire to the students.

The teachers in the four divisions were numbered for utilization in an APL program. The APL terminal has a random generator which will select as many numbers randomly from a range which the operator sets in the program. This range was the number of teachers in each division. Thus, the program randomly selected two teachers from each division. A similar system was used by the teachers to select a class. However, most teachers in the summer session have only two classes, so selection was simplified.

Procedure for Treating Data. Because of the dichotomous results in the frequencies of the first eight questions, a reclassification was

necessary. The eight questions were combined into four groups on the basis of similarity and relative ideas.

| | |
|---|---|
| Group 1<br>(Question 1) | Nominal Term<br>Authoritative teacher |
| Group 2<br>(Question 2 & 7) | Student-Center teacher -<br>A teacher who sees the student's uniqueness and helps him express his inner feelings. |
| Group 3<br>(Questions 3,4,5) | Social Conscious Teacher -<br>A teacher who sees social problems and promotes intellectual social problem solving. |
| Group 4<br>(Questions 6,8) | Career Counselor - Teacher -<br>Teacher who gives purpose to student's life and teaches practical skills for livelihood. |

After classifying the questions into groups, items (9) through (32) were summated for each of the four groups. That number was run as a parameter measure on an analysis of variance for the group.

Hypothesis: The four types of perceptions as catalogued in each of the groups do have signficantly different ratings for the teachers.

Null Hypothesis: The four types do not have signficantly different ratings for the teachers.

The data was treated with various computer programs. A multivariate Analysis of Variance distributed by Clyde Computing Service,

Box 166, Coconut Grove Station, Miami, Florida, 33133, was utilized. Also used was a program on internal reliability computed to produce Cronback's Coefficient a. A chi square was also run on items (9) through (32).

Limitation. The time limitation of completing this study, of course, produced some difficulties. As data was being collected the frequency distribution of the first eight questions were not realized until primary runs were made on the computer. An attempt to acquire a larger sample for better distribution was blocked due to the semester ending. The teachers, during the close of the final weeks of the summer session, were reluctant to take the forty minutes to complete the questionnaire in their classes, plus the fact that the students' attitudes at the end of a six week session are not favorable. However, the similarities of the groups formed are very good and allowed for a much stronger statistical instrument , an analysis of variance , to be used. (The sample was compiled in the winter semester.)

## RESULTS

Characteristics of the Population. Question 33 to 44 were Part III of the questionnaire. These questions provided more information on a description of the sample. Fifty-five percent of the students claimed a cumulative grade point average of 3 through 4, and thirty-six percent from a 2 to a 2.99. Thirty-six percent were earning an "A" in the course and seventy percent liked the subject they were presently taking.

Although school policy frowns upon working full time and going to school full time, a majority of the students do not keep this policy. Fifty-nine percent of the sample were taking at least fifteen semester hours of classes and fifty-eight percent were working more than twenty five hours/week. Thirty five percent were employed more than thirty five hours. A more complete summary is given in Appendix C.

Significance of Data

Analysis of Variance. The analysis of variance on scaled scores had one factor and one variable, but at four levels. Factor G as used in the table following are the four groups.

Table 1. Means and Standard Deviations

| Factor | Number | | Score |
|---|---|---|---|
| G | | | |
| 1 | 57 | | |
| | | M | 47.211 |
| | | SD | 15.922 |
| 2 | 36 | | |
| | | M | 43.028 |
| | | SD | 15.585 |
| 3 | 19 | | |
| | | M | 45.211 |
| | | SD | 14.497 |
| 4 | 36 | | |
| | | M | 50.500 |
| | | SD | 17.928 |

The analysis of variance of the four groups and the scaled scores generated the following results:

Table 2. Analysis of Variance

| Source | SS | DF | MS | F | P less than |
|---|---|---|---|---|---|
| Within cells | 37730.582 | 144 | 262.018 | | |
| G | 1062.120 | 3 | 354.040 | 1.351 | 0.260 |

According to the design of the above run (Edwards, 1972) the p value at a 0.260 level was not significant. Usually this result is considered too small a p value to be significant. Consequently, the above results do not allow us to reject our null hypothesis; that is, the measurable differences which were found to occur in the differences in perception of teaching as associated with differences in student evaluation of teachers could be attributed largely to mere chance fluctuations within our total population. The actual values were: 47.211 $\pm$ 15.922, Group I; 43.028 $\pm$ 15.585, Group II; 45.211 $\pm$ 14.497, Group III; and 50.500 $\pm$ 17.928 for Group IV.

Internal Reliability. The internal reliability computed from ANOVA was run on questions (9) through (32) for each of the groups. That run provided .909 internal reliability for Group I, 0.934 for Group II, 1905 for Group III, and .941 for Group IV. Thus, the internal reliability by groups was at an acceptable level.

## CONCLUSIONS AND RECOMMENDATIONS

It would appear on the basis of the research reviewed that an instructional rating system when properly developed can provide reliable and valid information about teaching. Research findings suggest that the criteria used by students in their ratings of instructors had much more to do with course objectives and consumer satisfaction than with entertainment value. Such attributes as preparedness, clarity, and stimulating of students' intellectual curiosity were typically mentioned

by students in describing their best instructor. Correlation for teacher personality characteristics were not fruitful except for the repeating characteristic that continued to appear in the research -- the characteristic of warmth toward students. Other correlations of student ratings which wer noted were: "majors" tended to rate courses more highly than "non-majors" in some cases; students required to take a course sometimes rated it lower than those for whom it was an elective; and experience and higher ranking instructors usually received higher ratings than did their less ranked or experienced colleagues. There was also some evidence that feedback in the form of student ratings may improve the teacher's performance. And last, but not least, the value system or perception of teaching a student has may influence the evaluation.

The results of this research confirms that the null-hypothesis may not be rejected in the population we studied herein. That is, students perception as associated with differences in student evaluation does not differ significantly from that which could be attributed to mere chance fluctuations within our total population. Because of the nearness of the end of the semester the resulting grouping may have weakened the instrument design of the selection of the course expectation. However, on a forced response of just the form groups, little difference in frequency distribution would probably have been observed if originally written in this fashion. This research indicates that further study would have to be done along the lines of students' perception influencing evaluation. Also, it is revealed here that students in the sample are oriented toward practical aspects of education as indicated by the number

who picked group four. The authoritative teacher still is preferred also.

If an evaluation system is to be adopted, it is recommended that a form which attempts to eliminate the students' value system be adopted. Frey (1974) claims to have this developed. If an evaluation system is to be adopted, it should be stressed that means or normals of different courses be kept for courses which are considered to be more difficult. Research indicated this is a preferred method. This helps to answer the questions "Which teachers are most effective?" and "For which objectives?" With such additional specification, student evaluation may provide useful evidence of teaching effectiveness.

"What is to be done?" This answer is discussed in Appendix A. page 25.

APPENDIX A

## SHOULD A STUDENT EVALUATION OF TEACHING BE ADOPTED IN THE COMMUNITY COLLEGE?

My main concern is to present rationales that will assure faculty that student evaluation would be an effective instrument to adopt at their college. A logical approach is to answer or discuss the objections which faculty have presented to me in the literature and research, and associates from different universities and present them with my comments.

1. The effectiveness of instruction cannot be judged until years after the fact.

The professor who is hated in undergraduate days yet revered in middle life "as the best damn teacher I had" has become trite. Although more study is needed, there is evidence that evaluations of teachers change little with time. Drucker and Remmers (1951) give evidence and more recently Centra (1974) provided even greater proof that there is no significant difference in student and alumni ratings.

2. Student evaluation of teaching is an invasion of academic freedom.

Academic freedom, as interpreted by most voters today, should not be interpreted as granting a professor immunity from orderly and responsible appraisal of his competence in so major an area of his professional activity as teaching. The judgement of academic performance by professional standards does not constitute interference with free inquiry and expression. Indeed, it may well be that responsible evaluation by professional standards of our devising is the only way that we can protect ourselves from more state or federal interference or influence.

3. The regular use of student evaluation of teaching would constitute a very big change in academic procedures some colleagues state.

This may be true that many would view the student evaluation as a big change. They have lived with the present system and have survived within it. Any change is seen as a threat by some of them. Another view is that the changes would only help us to do better what we already are trying to do -- give recognition to quality of instruction when recommending promotion and merit increases -- and hence, it would not be such a big change. No sudden big change needs to happen. Prudence might counsel starting with a pilot project involving selected departments.

4. How can we account for racial, ethnic, and sex biases of students in evaluating teachers?

This problem may become more acute if the Civil Rights Act is successful in placing women and other minority group members in high prestige positions if they are not qualified. Many students find women and minority groups a threat and may find it difficult to evaluate them. The only scholar who has entered into thought on this problem is A. T. Sharon (1970). He feels that in the structure of the evaluation you can eliminate some bias.

5. The most recent article to ignite thought is Irene R. Kerman's "Student Evaluation Re-evaluated (1975)".

Kierman stated that:

> Student evaluation of faculty falls under the definition of "test" as defined by the American and Psychological Association and Equal Employment Opportunities Commission Guidelines concerning appointment, reappointments, and promotion of faculty members.

Kierman claims the teacher evaluation tests have not been validated especially for minority groups. Thus she claims they can challenge their use.

This is not true. Many studies as McKeachie (1971) and Costin (1971) have developed valid testing procedures. The Educational Testing Service of New Jersey has developed a system called "SIR", Student Instructional Report.

6. Regularized ratings of teaching would constitute another instance of the dehumanizing of people by describing them with numbers, a trend which is deplored by students.

Such numbers, like many others, are not primary sources of dehumanization. Numbers are merely symptoms of the estrangement that comes with too many people. Stamping out numbers will not cure the malady deplored by students. At a college where ten professors teach one hundred students, numerical ratings of teaching are not useful, everyone carries sufficient evaluation in his head. But, at a large university, it is more dehumanizing to fail to recognize and reward effective teaching adequately.

7. Prior ratings would prejudice subsequent evaluations.

This is called the "halo effect." If ratings are not published, this is not a consideration. Hal effect is not a serious threat. Professors make great show of their research or national recognition to committees, why not include student ratings. Besides, professors are a frequent topic of conversation among students and they are continually being rated.

I do not fancy my brief comments have disposed of the above arguments against student evaluation of teaching. Some of the objections are well

taken, and the entire subject of evaluation presents a complex pattern of considerations combining white, black, and shades of gray. I do believe, however, that the weight of reasoned evidence justifies giving student evaluation of teaching a thorougn trial.

What is to be done? I believe it is desirable to: (1) establish a regular use of student evaluation of teaching for the benefit of the instructors, students, and advancement procedures; (2) gain the cooperation of the entire faculty to secure the ratings and establish the purposes, objectives, and uses; (3) norms be calculated and a counselor be established to interpret the results if requested ; (4) student evaluation of teaching be used to supplement, but not substitute for other valid kinds of evaluation. Similar conclusions have been reached by many persons who have studied the evaluation of teaching as Kenneth E. Eble, Director of the Project to Improve College Teaching, (Eble, 1974)

Numerous problems remain on which informed and reasonable persons have differing opinions: How will the procedures adopted give the desired emphasis to the two primary objectives of: (1) contributing indirectly to the betterment of teaching by improving the evaulation of teaching; and (2) contributing directly to the betterment of teaching by helping individual instructors to improve. Surely the academic community could find adequate solution to these pr blems so that a flexible program could be initiated. This is, it could if it would -- and I am conficent that in time it will, as soon as leadership appears.

STUDENT EVALUATION OF THE TEACHING OF MIAMI-DADE COURSES

The purpose of this evaluation form is to furnish a basis for the continuous improvement of instruction. When these evaluation forms are used at the end of this term, they will not be looked at by the instructor until final grades have been filed. Your response is anonymous. DO NOT FILL IN YOUR NAME. After you have completed the 44 questions, hand in your answer card and questions. Mark your choice of answers on the card provided. After you have read ALL of the eight following questions, pick out the ONE statement which MOST ACCURATELY reflects your concept of what teaching and learning mean, and mark the (1) STRONGLY AGREE space on your IBM card. Use the number "1" space ONLY ONE TIME for the first eight statements. Proceed to rank the remainder (7 questions) between choices "2" and "5". For the first eight questions use the following scale: DO NOT WRITE ON THIS FORM.

1. Strongly agree
2. Agree
3. Not sure whether I agree or disagree
4. Disagree
5. Strongly disagree

1. I like a teacher best who knows his field and authoritatively declares that information to the class.

2. I like a teacher best who develops a climate in which I may express my inner feelings and clarify them to myself.

3. I like a teacher who motivates problem-solving activities by pointing out the gap between what ought to be done in today's society about social problems and what is being done about them.

4. I like a teacher who helps individuals resolve the conflicts in society, for example, justice versus injustice, equality versus inequality.

5. I like a teacher who helps me solve tough, complex intellectual problems.

6. I like a teacher whose course gives me direction and purpose for my life.

7. I like a teacher whose classes emphasize the student's uniqueness and individuality.

8. I like a teacher whose primary purpose is to teach me skills I can utilize in earning a living.

MARK THE SCALE PROVIDED ON THE REMAINDER OF THE QUESTIONS.

| | 1<br>Excellent | 2 | 3<br>Fair | 4 | 5<br>Very Poor |
|---|---|---|---|---|---|
| 9. How do you rate him(her) as a teacher? | _____ | _____ | _____ | _____ | _____ |
| 10. How is his knowledge of his subject? | _____ | _____ | _____ | _____ | _____ |

| | | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| 11. | How is the organization of his course? | Excellent | | Fair | | Very Poor |
| 12. | Are his explanations in class clear? | Always | | Usually | | Seldom |
| 13. | How are his assignments: | Clear | | Usually clear | | Confusing |
| 14. | How enthusiastic is he in class? | Very | | Usually | | Lacks enthusiasm |
| 15. | How is his ability to express his thoughts: | Excellent | | Usually has no difficulty | | Has a great deal of difficulty |
| 16. | How is his sense of stressing important matters? | Excellent | | Fair | | Very poor |
| 17. | How is his ability to stick to the subject? | Always sticks to subject | | Rarely digresses | | Gets completely off subject |
| 18. | Is opportunity given for students to participate in class? | Often | | Occasionally | | Rarely |
| 19. | How is his ability to inspire pupils? | Excellent | | Fair | | Very poor |
| 20. | How much does he encourage the student to think for himself? | A great deal | | Some | | Not at all |
| 21. | How are his tests as to coverage of material? | Very reasonable | | Usually reasonable | | Unreasonable |

| | | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| 22. | How is his grading as to fairness? | Very fair | | Usually fair | | Very unfair |
| 23. | What is his attitude toward difference of opinions on controversial questions? | Very tolerant | | Fairly tolerant | | Very intolerant |
| 24. | How is his attitude as an instructor toward student? | Very understanding & patient | | Usually patient & responsive | | Impatient & indifferent |
| 25. | How is his sense of humor? | Exceptionally good | | Fair | | Very little |
| 26. | How often have you seen cheating in his class? | Never | | Sometimes | | Very often |
| 27. | How is his personal appearance? | Very neat | | Usually neat | | Very careless |
| 28. | How are his classroom manners? | Very courteous | | Usually courteous | | Very discourteous |
| 29. | How is his speech? | Easily understood & pleasant | | Sometimes inaudible & indistinct | | Very difficult to listen to |
| 30. | How about his mannerisms? | Usually attractive classroom manners | | Moderately free from distracting mannerisms | | Frequent annoying mannerisms |
| 31. | Overall rating of course. | An outstanding course | | A reasonably good course | | A very poor course |
| 32. | If given the opportunity I would take another course with this instructor. | Yes | | Maybe | | No |

33. Your cumulative grade point average is:
1. 3.0 through 4.0 2. 2.0 through 2.99 3. 1.0 through 1.99
4. 0.99 or less

34. The grade you are presently earning in this course is:
1. A 2. B 3. C 4. D 5. F

35. If employed, how many hours per week?
1. less than 16 2. 16 through 24 3. 25 through 34
4. 35 through 40 5. over 40

36. How many credit hours are your presently taking?
1. Under 12 2. 12 through 15 3. 16 through 18 4. Over 18

37. Sex: 1. Female 2. Male

38. Why are you taking this course:
1. A requirement for your major. 2. An elective for your major
3. A requirement outside your major 4. An elective outside your major.

39. To what extent have you applied yourself in this course?
1. To the maximum 2. Very much 3. Average 4. Very little
5. Not at all

40. Do you like this subject?
1. Yes 2. No

41. Do you feel you had enough background to take this course?
1. Yes 2. No

42. Do you attend class regularly?
1. Yes 2. No

43. Are you:
1. Black American 2. White American other than (3) 3. Spanish surnamed
4. Foreign student other than (3). 5. Other

44. This course is a course in:
1. Science or Math 2. Social Science 3. English or Language
4. Other technical fields (TVS) 5. Other

APPENDIX C

*

33. Your cumulative grade point average is:
    1. 3.0 through 4.0 (55%)  2. 2.0 through 2.99 (36%)
    3. 1.0 through 1.99 (3%)  4. 0.99 or less (1%)  5. 1

34. The grade you are presently earning in this course is:
    1. A 36%  2. B 32%  3. C 23%  4. D 3%  5. F 2%

35. If employed, how many hours per week?
    1. less than 16 (20%)  2. 16 through 24 (23%)
    3. 25 through 34 (13%)  4. 35 through 40 (24%)  5. over 40 (11%)

36. How many credit hours are you presently taking?
    1. under 12 (38%)  2. 12 through 15 (33%)  3. 16 through 18 (18%)
    4. over 18 (8%)

37. Sex: 1. Female 39%  2. Male 55%

38. Why are you taking this course?
    1. A requirement for your major (63%)
    2. An elective for your major (13%)
    3. A requirement outside your major (12%)
    4. An elective outside your major. (7%)

39. To what extent have you applied yourself in this course?
    1. To the maximum (24%)  2. Very much (38%)
    3. Average (28%)  4. Very little (6%)  5. Not at all (2%)

40. Do you like this subject?
    1. yes (70%)  2. No (21%)

41. Do you feel you had enough background to take this course?
    1. Yes (75%)  2. No (21%)

42. Do you attend class regularly?
    1. yes (80%)  2. No (17%)

43. Are you:
    1. Black American (14%)  2. White American other than (3) (46%)
    3. Spanish Surnamed (23%)  4. Foreign student other than (3) (6%)
    5. Other (6%)

44. This course is a course in:
    1. Science or Math (55%)  2. Social Science (16%)
    3. English or Language (9%)  4. Other technical fields (TVS) (3%)
    5. Other (11%)

* The small percentage that is missing in some of the statements were a result of blanks. Each question had no more than two or three blanks.

BIBLIOGRAPHY


1. Brower, F.B. Personality Assessment of College and University Faculty: Implications for the Junior College. Washington, D.C.: American Association for Junior Colleges, 1968.

2. Centra, John R. Strategies for Improving College Teaching. Washingon, D.C.: American Association of Higher Education, December 1972.

3. Centra, John R. The Utility of Student Ratings for Instructional Improvement. PR 72-16, Princeton; Educational Testing Service, August 1972b.

4. Centra, John R. "Do Student Ratings of Teachers Improve Instruction?" Change. Vol. 5 #3, April 1973a, pp 12-13.

5. Centra, John R. "The Relationship between Student and Alumni Ratings of Teachers" Educational & Psychological Measurement. Summer 1974. Erie Clearing House, Report 9.

6. Cohen, J. and L. G. Humphreys. Memorandum to Faculty. University of Illinois, Department of Psychology, 1960. (mimeographed)

7. Cohen, Arthur M. and Florence B. Brawer. "Measuring Faculty Performance". Washington, D.C.: ERIC Clearing House for Junior Colleges, 1969.

8. Costin, Frank, W. T. Greenough, and R. J. Menzes. "Student Ratings of College Teaching: Reliability, Validity, and Usefulness". Review of Educational Research 41, no. 5. (December 1971)

9. Cronback, L. J. and L. Furby. "How We Should Measure 'Change' -- or Should We?" Psychological Bulletin, 1970, 74, 68-80.

10. Downie, N.W. "Student Evaluation of Faculty", Journal of Higher Education, 1952,23, 495-496.

11. Drucker, A. J. and H. H. Remmers. "Do Alumni and Students Differ in Their Attitudes Toward Instructors?" Journal of Educational Psychology, XLII (March, 1951) pp. 129-43.

12. Eble, Kenneth. Career Development of the Effective College Teacher. Washington, D.C.: American Association of University Professors, November 1971.

13. Eble, Kenneth. The Recognition and Evaluation of Teaching American Association of University Professors. February 1974.

14. Edwards, Allen I. Experimental Design (4th Edition) New York: Holt, Rinehard, and Winston 1972.

15. Elliott, D. H. Characteristics and Relationship of Various Criteria of College & University Teaching. Purdue University Studies in Higher Education, 1950.

16. French, G. M. "College Students' Concept of Effective Teaching Determined By An Analysis of Teacher Ratings". Dissertation Abstract, 17:1380-1381.

17. Frey, Peter W. "Student Rating of Teaching: Validity of Several Ratings Factors". Science Vol 182, No. 4107, October 5, 1973, pp. 83-85.

18. Frey, Peter W. "The Ongoing Debate: Student Evaluation of Teaching." Change Vol. 6, No. 1, February 1974.

19. Gage, N. J. "The Appraisal of College Teaching" Journal of Higher Education, 1961, 32. pp. 17-22.

20. Guthrie, E. R. The Evaluation of Teaching: A Progress Report. Seattle: University of Washington, 1954.

21. Hildebrand, Milton. "How to Recommend Promotion for a Mediocre Teacher Without Actually Lying" Journal of Higher Education, 43, 1972, 44-62.

22. Kierman, Irene R. "Student Evaluations, Re-evaluationed" Community and Junior College Journal, Vol. 7, No. 7, April 1975.

23. Lovell, G. D. and C. F. Haner. "Force-choice Applied to College Faculty Rating" Educational and Psychological Measurement, 1955, 15,291-304.

24. Lothrop, R. G. and C. Richmond. College Students' Evaluation of Courses and Instructors. Chico State College, California, 1967.

25. McDaniel, D. and J. F. Feldhusen. Relationship Between Faculty Ratings and Indexes of Service and Scholarship. Proceedings of the 78th Annual Convention of the American Psychological Association, 1970,5, 619-620

26. McKeachie, W. J. Teaching Tips: A Guidebook for the Beginning College Teacher. (6th Edition) Lexington, Mass: D. C. Health, 1969.

27. McKeachie, W. J., Y Linn, and W. Mann. Student Ratings of Teacher Effectiveness: Validity Studies. American Educational Research Journal, 1971, 8, 435-45.

28. Miller, M. T. "Instructor Attitudes Toward and Use of Student Ratings of Teachers" Journal of Educational Psychology 62(1971): 235-39.

29. Perry, R. D. Criteria of Effective Teaching in an Institution of Higher Education. Office of Institutional Research. University of Toledo, 1967.

30. Quick, F. A. and A. D. Wolfe. "The Ideal Professor", Improving College And University Teaching 13, no. 3 (1965)

31. Roden, Miriam. "Can Student Evaluate Good Teaching". Change, Vol. 7, No. 6, Summer 1973.

32. Sharon, A. T. "Elementary Bias from Student Rating of College Instructors" Journal of Applied Psychology, 54:278-81, 1970.

33. Sullivan, Arthur. Meeting presented at American Psychological Association. Science.

34. Walker, B. D. An Investigation of Selected Variables Relative to the Manner in Which a Population of Junior College Students Evaluate Their Teachers. Dissertation Abstracts, 1969, 57, 21-25.

UNIVERSITY OF CALIF.
LOS ANGELES

MAR 19 1976

CLEARINGHOUSE FOR
JUNIOR COLLEGES